

Relation of Surgical Volume to Outcome in Eight Common Operations

Results From the VA National Surgical Quality Improvement Program

Shukri F. Khuri, MD,* Jennifer Daley, MD,† William Henderson, PhD,‡ Kwan Hur, PhD,‡ Monir Hossain, MS,*** David Soybel, MD,* Kenneth W. Kizer, MD, MPH,§ J. Bradley Aust, MD,|| Richard H. Bell, Jr., MD,¶ Vernon Chong, MD,# John Demakis, MD,‡ Peter J. Fabri, MD,** James O. Gibbs, PhD,‡ Frederick Grover, MD,†† Karl Hammermeister, MD,†† Gerald McDonald, MD, MS,‡‡ Edward Passaro, Jr., MD,§§ Lloyd Phillips, MD,||| Frank Scamman, MD,¶¶ Jeannette Spencer, RN, MS, CS,*** John F. Strempel, MD, MS,## and the participants in the VA National Surgical Quality Improvement Program

*From *Brockton/West Roxbury VA Medical Center, West Roxbury, and Harvard Medical School and Brigham and Women's Hospital, Boston, MA; †Brockton/West Roxbury VA Medical Center, West Roxbury, and Harvard Medical School and the Division of General Medicine and Primary Care, Beth Israel Deaconess Medical Center, Boston, MA; ‡Hines VA Cooperative Studies Program Coordinating Center, Hines, IL; §Undersecretary for Health, Department of Veterans Affairs Headquarters, Washington DC; ||Department of Surgery, University of Texas Health Science Center, San Antonio, TX; ¶VA Puget Health Care System and University of Washington, Seattle, WA; #Veterans Integrated Service Network 17, Grand Prairie, TX; **Tampa VA Medical Center and University of South Florida College of Medicine, Tampa, FL; ††Denver VA Medical Center and University of Colorado Health Sciences Center, Denver, CO; ‡‡Department of Surgery, Department of Veterans Affairs Headquarters, Washington DC; §§West Los Angeles VA Medical Center, West Los Angeles, CA, and University of California Los Angeles, Los Angeles, CA; |||Veterans Integrated Service Network 11, Ann Arbor, MI; ¶¶Department of Anesthesia, National Anesthesia Services, Iowa City, IA; ##Pittsburgh VA Medical Center and University of Pittsburgh School of Medicine, Pittsburgh, PA; ***Brockton/West Roxbury VA Medical Center, West Roxbury, MA*

Objective

To examine, in the Veterans Health Administration (VHA), the relation between surgical volume and outcome in eight commonly performed operations of intermediate complexity.

Summary Background Data

In multihospital health care systems such as VHA, consideration is often given to closing low-volume surgical services, with the assumption that better surgical outcomes are achieved in hospitals with larger surgical volumes. Literature data to support this assumption in intermediate-complexity operations are either limited or controversial.

Methods

The VHA National Surgical Quality Improvement Program data on nonruptured abdominal aortic aneurysmectomy, vascular infrainguinal reconstruction, carotid endarterectomy (CEA), lung lobectomy/pneumonectomy, open and laparoscopic cholecystectomy, partial colectomy, and total hip arthroplasty

were used. Pearson correlation, analysis of variance, mixed effects hierarchical logistic regression, and automatic interaction detection analysis were used to assess the association of annual procedure/specialty volume with risk-adjusted 30-day death (and stroke in CEA).

Results

Eight major surgical procedures (68,631 operations) were analyzed. No statistically significant associations between procedure or specialty volume and 30-day mortality rate (or 30-day stroke rate in CEA) were found.

Conclusions

In VHA hospitals, the procedure and surgical specialty volume in eight prevalent operations of intermediate complexity are not associated with risk-adjusted 30-day mortality rate from these operations, or with the risk-adjusted 30-day stroke rate from CEA. Volume of surgery in these operations should not be used as a surrogate for quality of surgical care.

The Veterans Health Administration (VHA) operates the largest fully integrated health care system in the United States.¹ Within this system, 123 hospitals perform major surgery in surgical services varying widely in size and in the volume of operations performed annually. To reduce the cost and improve the quality of surgical care in multihospital health care systems such as VHA, consideration is often given to closing low-volume surgical services, with the assumption that better surgical outcomes are achieved in hospitals with larger surgical volumes. This assumption is based on studies that show an inverse correlation between the volume of surgery and mortality or morbidity rates in cardiac surgery² and in complex noncardiac operations such as abdominal aortic aneurysmectomy,^{3,4} carotid endarterectomy (CEA),^{4,5-12} and pancreaticoduodenectomy.¹³⁻¹⁵ Most of the studies addressing the relation between surgical volume and outcome, however, are limited by their retrospective nature, the administrative databases on which they were based, the selection of the institutions included in the analyses, and the failure to adjust for differences in patient preoperative risk factors in the assessment of deaths and other adverse postoperative outcomes.^{16,17} In addition, selection bias—the transfer of less sick surgical candidates to high-volume institutions for major surgery—has not been systematically evaluated as a possible source for the observed lower mortality rate at high-volume hospitals.

This study was undertaken to examine the relation between surgical volume and outcome in eight commonly performed operations in VHA, using the FY91–FY97 database of the VA National Surgical Quality Improvement Program (NSQIP). This database contains preoperative patient risk factors, operative data, and 30-day outcome information on all major surgeries performed in VHA, collected prospectively by a dedicated nurse reviewer at each medical center.^{18,19} The models developed by the NSQIP for risk adjustment of 30-day postoperative mortality and morbidity rates have been validated and shown to reflect the quality of surgical care.²⁰⁻²²

METHODS

Overview of the NSQIP

The NSQIP is an ongoing quality management initiative that applies the methodology developed and validated by

the National VA Surgical Risk Study (NVSRS) to all of the Veterans Affairs medical centers (VAMCs) that perform major surgery. Both the NSQIP and the NVSRS have been described in detail in previous publications.^{18-20,23,24} A brief summary is outlined below.

Participating Centers

Between October 1991 and December 1993, 44 VAMCs then performing cardiac and noncardiac surgery contributed preoperative patient risk and operative and postoperative outcome data about major operations to the NVSRS. Since the inception of the NSQIP in 1994 and the expansion of data collection and reporting to all VAMCs performing major surgery, nine VAMCs have stopped performing major surgery, bringing the number of participating VAMCs as of January 1999 to 123.

Common Operations Selected for Volume/Outcome Analysis

We selected eight commonly performed operations for this volume/outcome analysis. The CPT-4 codes for the selected operations are listed in Table 1. In vascular surgery, we selected nonruptured abdominal aortic aneurysm, vascular infrainguinal reconstruction, and CEA. In orthopedic surgery, we selected total hip arthroplasty; in general surgery, partial colectomy, open cholecystectomy, and laparoscopic cholecystectomy; and in noncardiac thoracic surgery, lung lobectomy and pneumonectomy. Laparoscopic cholecystectomies that were converted to open cholecystectomies during surgery were identified as having CPT-4 codes for both laparoscopic and open operations and were treated in the analysis as open operations. Each of these operations is defined as major surgery and was performed under general, spinal, or epidural anesthesia, with the exception of CEA, which may have been performed under local anesthesia. Carotid endarterectomies may have been performed by surgeons who identified themselves as general, vascular, or neurosurgeons at the time of the operation.

Data Collection

A trained surgical clinical nurse reviewer collects and verifies 65 preoperative patient characteristics, 11 intraoperative variables (including up to three CPT-4 codes identifying each operation), and 23 outcomes, including mortality status, and 20 uniformly defined postoperative adverse events at 30 days after surgery. These data are verified by the chief of surgery and transmitted electronically to a central data repository at the Hines VA Cooperative Studies Program Coordinating Center. Detailed descriptions of nurse reviewer training and supervision, data collection protocols, and data verification and cleaning procedures may be found in other publications.^{18,20,23}

Presented at the 119th Annual Meeting of the American Surgical Association, April 15–17, 1999, Hyatt Regency Hotel, San Diego, California. The Department of Veterans Affairs National Surgical Quality Improvement Program is funded by the Veterans Health Administration of the Department of Veterans Affairs.

Dr. Daley was a Senior Research Associate in the Career Development Award Program of the Veterans Affairs Health Services Research and Development Service during the period of this report.

Correspondence: Shukri F. Khuri, MD, Veterans Affairs Medical Center, 1400 VFW Parkway, West Roxbury, MA 02132

Accepted for publication April 1999.

Table 1. COMMON PROCEDURAL TERMINOLOGY-4 (CPT-4) CODES USED IN VOLUME-OUTCOME ANALYSIS

Procedure	CPT-4 Codes
Abdominal aortic aneurysmectomy	35081
Infrainguinal vascular reconstruction	Bypass graft–Vein 35521, 35533, 35546, 35548, 35549, 35551, 35556, 35558, 35565, 35566, 35571 Bypass graft–In-situ vein 35582, 35583, 35587 Bypass graft–Other than vein 35621, 35623, 35646, 35651, 35654, 35656, 35661, 35665, 35666, 35671
Carotid endarterectomy	35301
Lobectomy/pneumonectomy	Lobectomy 32480, 32485, 32490 Pneumonectomy 32440, 32445
Open cholecystectomy	47600, 47605, 47610
Laparoscopic cholecystectomy	56340, 56341, 56342, 49310, 49311
Colectomy	44140, 44141, 44143, 44144, 44145, 44146, 44147, 44150, 44151, 44152, 44153, 44155, 44156, 44160
Total hip arthroplasty	27130, 27131, 27132, 27134

Definition of Volume

Because each of the 123 VAMCs currently participating in the NSQIP may have contributed data to the NSQIP for a different number of months, we defined the volume of each operation to be the total number of cases of each operation in the database divided by the number of months that the VAMC had contributed to the NVASRS and NSQIP databases and multiplied by 12 to arrive at an annual VAMC volume for each operation. Volume variables were computed both at the individual operation level and at the specialty level.

Definition of Outcomes

Thirty-day postoperative mortality was obtained from all patients having any one of the eight common operations. All deaths were verified against the VHA Beneficiary Identification and Records Locator Subsystem death records. For CEA, we also determined the presence of postoperative stroke at 30 days, defined as the development of an embolic, thrombotic, or hemorrhagic vascular accident or stroke with motor, sensory, or cognitive dysfunction that persists for ≥ 24 hours.

Statistical Analysis

A logistic regression analysis was performed for each of the eight operations, with 30-day mortality as the dependent variable and the patient risk factors as the independent variables. For CEA with stroke as the outcome, the logistic regression analysis used stroke as the dependent variable. The c-index, a measure of predictive validity of the model, is the proportion of all possible pairs of concordant cases (dead/alive or stroke/no stroke cases) for which the proba-

bility of the event calculated from the logistic regression equation is higher for the case with the event than for the case without the event.

After the nine logistic regression models were developed, the models were used to calculate the predicted probability of an event for each case. These predicted probabilities were summed to arrive at the expected number of deaths (or strokes) at each VAMC for each operation. The ratio of observed number of deaths (or strokes) to expected number of deaths (or strokes)—the O/E ratio—is a measure of the risk-adjusted outcome for an operation at a particular VAMC. An O/E ratio greater than (less than) one is an indication that the VAMC is experiencing more (fewer) events than would have been expected after adjustment for the burden of illness in that VAMC patient population.

Pearson correlation coefficients were computed correlating O/E ratios and operation or specialty volumes across all VAMCs for each operation. A positive (negative) correlation indicates that as volume increases, risk-adjusted outcomes become worse (better). Probability values are given for the test of the correlation coefficient equaling zero *versus* not zero.

A mixed effects hierarchical logistic regression model was also used to assess the volume/outcome relation for each operation. In the first level of modeling, the patient is the unit of analysis and the model accounts for the impact of patient risk factors on the outcome of death in the eight operations and stroke in CEA. In the second level, the association of annual hospital volume for the operation and specialty with risk-adjusted outcome (death, stroke for CEA) is assessed, and the hospital is the unit of analysis. Two mixed hierarchical logistic regression models were created for each operation. In the second level of each model, the first model uses annual operation volumes as the

Table 2. PATIENT DEMOGRAPHICS AND 30-DAY MORTALITY BY OPERATION

Operation	No. of Cases	Age (years)		Sex		Race		Emergent Status		30-Day Mortality		Excluded and Nonassessed Cases**		
		Mean \pm SD	Range	No. Male	%	No. White	%	No. Emergent	%	No. of Deaths	%	Number	No. of Deaths	%
Abdominal aortic aneurysmectomy	3767	68.8 \pm 6.9	36–98	3745	99.4	3288	88.7	227	6.0	177	4.7	281	9	3.2
Infrainguinal vascular reconstruction	12535	64.5 \pm 9.2	22–97	12424	99.1	9510	77.1	747	6.0	383	3.1	1185	45	3.8
Carotid endarterectomy*	10173	67.2 \pm 7.8	27–100	10021	98.5	8962	89.4	255	2.5	123	1.2	1710	30	1.8
Lobectomy/pneumonectomy	4890	64.7 \pm 9.0	21–91	4828	98.7	3858	80.3	31	0.6	267	5.5	358	17	4.7
Open cholecystectomy	7113	62.6 \pm 12.2	20–102	6874	96.6	5375	76.9	1221	17.2	201	2.8	461	18	3.9
Laparoscopic cholecystectomy	8602	57.9 \pm 37.4	16–97	7818	90.9	6629	78.6	214	2.5	46	0.5	763	6	0.8
Colectomy	13310	66.4 \pm 11.1	19–100	13056	98.1	9785	75.0	2473	18.6	922	6.9	990	73	7.3
Total hip arthroplasty	8241	63.2 \pm 11.8	20–103	8028	97.4	6354	78.2	69	0.8	84	1.0	727	6	0.8
Total	68631	64.3 \pm 11.0	16–103	66794	97.3	53761	78.3	5237	7.6	2203	3.2	6475	204	3.2

* Stroke as a 30-day outcome occurred in 212 patients (2.1%).

** No statistically significant difference ($p < 0.05$) in 30-day mortality rate between excluded/nonassessed cases and those in the analytic data set.

independent variable; the second model uses the annual volume of major operations performed in the surgical specialty as the independent variable.

To determine whether there was a threshold volume effect below which outcomes worsen, three analyses were performed:

1. Indicator variables for the four quartiles of procedure and specialty volumes were added to the mixed effects logistic regression model.
2. Analysis of variance was used to compare the mean O/E ratios across the four quartiles of procedure volume.
3. An automatic interaction detection statistical analysis^{25,26} was performed using a set of algorithms contained in the PC Group statistical software.²⁷ These algorithms searched for all possible cutpoints in the procedure and specialty volumes below which the outcomes worsened and attempted to identify the cutpoints that minimized the overall misclassification rate.

The initial models for CEA, which were developed based on all 101 hospitals in which this operation was performed, had low c-indices, indicating poor predictive validity. The c-indices improved, and the models became more predictive, when eight low-volume hospitals that had collected a truncated set of the preoperative variables were excluded from the analysis. Hence, data from 93 hospitals only were used in the CEA analyses.

Results were considered statistically significant at $p \leq 0.05$.

RESULTS

Patient and Hospital Characteristics

Eight major surgical procedures totaling 68,631 operations were analyzed. As shown in Table 2, the total volume of each operation ranged from 3767 abdominal aortic aneurysm repairs to 13,310 partial colectomies. The mean age ranged from 57.9 years in patients undergoing laparoscopic cholecystectomy to 68.8 years in patients undergoing abdominal aortic aneurysm repairs. Laparoscopic cholecystectomy cases had the highest percentage of women (9.1%); patients undergoing abdominal aortic aneurysm repairs had the lowest (0.6%). The percentage of whites ranged from 75% in patients undergoing partial colectomy to 89.4% in patients undergoing CEA. Only 0.6% of the pulmonary resection cases were emergent, in contrast to 18.6% of the partial colectomy cases. The observed 30-day mortality rate ranged from 0.5% in laparoscopic cholecystectomy to 6.9% in partial colectomy.

Table 2 also shows the patients in the NSQIP database who were excluded from the analysis. Reasons for excluding patients from risk assessment have been described elsewhere.^{18,23} Most of the exclusions in this study were because the clinical nurse reviewers in the respective hospitals were on annual leave at the time of the operation. As shown in Table 2, the mortality rate of the patients excluded from the analysis was not significantly different from the mortality rate of patients included in the analysis.

The number of hospitals performing these operations ranged from 125 hospitals performing partial colectomy to 93 hospitals performing CEA (Table 3). Infrainguinal vascular reconstruction, with a mean of 23.5 operations per

Table 3. SPECIALTY AND PROCEDURE VOLUME/HOSPITAL/YEAR

Operation	No. of Hospitals	Specialty Volume/Year		Procedure Volume/Year		% Performed by Resident
		Mean \pm SD	Range*	Mean \pm SD	Range*	
Abdominal aortic aneurysmectomy	107	89.7 \pm 53.8	1–240	6.9 \pm 5.7	0–32	76.6
Infrainguinal vascular reconstruction	107	89.7 \pm 53.8	2–240	23.5 \pm 14.8	1–90	76.1
Carotid endarterectomy	93	102.0 \pm 52.0	6–260	21.9 \pm 14.7	0–73	81.0
Lobectomy/pneumectomy	107	32.9 \pm 24.2	0–169	9.0 \pm 6.5	0–44	76.4
Open cholecystectomy	124	186.4 \pm 79.5	13–414	11.7 \pm 7.1	1–39	68.9
Laparoscopic cholecystectomy	123	187.5 \pm 78.9	13–414	15.1 \pm 8.0	0–44	68.4
Colectomy	125	185.5 \pm 80.6	13–414	21.9 \pm 11.5	0–52	69.6
Total hip arthroplasty	101	154.1 \pm 79.9	1–442	16.1 \pm 9.1	0–55	70.5

* 0 appears as a lower limit in hospitals that did not perform that procedure in the course of one or more years throughout the duration of the study.

hospital annually, was the most frequently performed operation; abdominal aortic aneurysm repair was the least frequently performed (mean 6.9 operations per hospital annually). In a relatively high percentage of operations (range 68.4% to 81.0%), a resident was listed as the primary surgeon, indicating that the operation was primarily performed by a resident in the presence of or with the assistance of a staff attending surgeon.

Risk Adjustment Models of 30-Day Outcomes

A total of nine models were constructed. For each of the eight operations, a logistic regression model was constructed that identified the significant preoperative patient risk factors predictive of 30-day mortality. In addition, a model was constructed for CEA identifying the preoperative patient risk factors predictive of postoperative stroke within 30 days. The preoperative patient characteristics that were significant in these models are listed in Table 4. The order of entry, which reflects the relative importance of each of these variables in each multivariable model, is also shown. Partial colectomy had the largest number of preoperative predictors; ASA class was the most important, followed by preoperative serum albumin level and emergency status. ASA class, age, and preoperative serum albumin level appeared in seven of the nine models.

The predictive validity of each model is indicated by the c-index,²⁸ which is also shown in Table 4. The partial colectomy model, with a c-index of 0.85, had the most predictive validity; the CEA models were the least predictive of both 30-day mortality and postoperative stroke (c-indices 0.72 and 0.64, respectively). Adequate predictive validity is usually indicated by a c-index of >0.70 .

Relation of Surgical Volume to Outcome

Pearson Correlation

The procedure volume and its respective specialty volume per hospital per year were each correlated separately to

their respective hospital-specific procedure risk-adjusted 30-day mortality rate (and 30-day stroke rate for CEA), expressed as the O/E ratio. The results of these analyses, shown in Table 5, indicated that there was no significant correlation between hospital volume (at both the procedure and the surgical specialty levels) and risk-adjusted 30-day mortality rate in all eight procedures. A weak correlation was observed between the specialty volume and risk-adjusted stroke rate after CEA. As specialty volume increased, risk-adjusted stroke rate improved. However, the correlation with procedure volume was not significant (see Table 5).

Figure 1, which illustrates these results, is a scattergram depicting each hospital's annual procedure volume plotted against the 30-day mortality O/E ratio for each of the eight operations and against the 30-day stroke O/E ratio for CEA. There was in general a wider variation in the O/E ratio between hospitals performing low volumes of surgery (on the left side of the graph) *versus* hospitals performing high volumes (on the right side of the graph). This is because the estimate of the O/E ratio tends to be less stable when sample sizes are small and is not necessarily a reflection of more variability in the quality of care at lower-volume hospitals.

Analysis of Variance Between Hospitals in Quartiles

To determine whether there was a volume threshold that might significantly affect outcome in the eight operations, hospitals were grouped in quartiles according to their procedure volume per year. An analysis of variance was then performed to determine whether there was a significant difference in O/E ratios between these quartiles. The results of this analysis are shown in Table 6. There were no significant interquartile differences in the O/E ratio in any of the operations, indicating that there was no volume threshold below which the risk-adjusted mortality rate (and the risk-adjusted stroke rate after CEA) was significantly increased. As shown in Figure 1, Table 6 also shows that the hospitals in the lower-volume quartiles exhibited larger standard deviations in the O/E ratio than the hospitals in the

Table 4. ORDER OF ENTRY OF PREOPERATIVE PREDICTOR VARIABLES AND C-INDICES OF EIGHT RISK ADJUSTMENT MODELS

Variable	Abdominal Aortic Aneurysmectomy	Infringuinal Vascular Reconstruction	Carotid Endarterectomy (Mortality)	Carotid Endarterectomy (Stroke)	Lobectomy/Pneumonectomy	Open Cholecystectomy	Laparoscopic Cholecystectomy	Colecotomy	Total Hip Arthroplasty
Emergency	1	2	1	3	4	3		3	
ASA class	2	1	6					1	3
WBC > 11.0 ($\times 10^3/\text{mm}^3$)	3	6							
BUN > 40 (mg/dl)	4	10			6	1	4	7	
Age (years)	5	3			3	4	3	5	4
Weight loss	6					8			
Albumin (gm/dl)	7	4			2	2	1	2	1
Dyspnea		9			5	5		4	2
SGOT > 40 (IU/ml)		5				9			5
Functional status			3		8	7	5		6
Dis. cancer								6	
DNR								8	
Bilirubin > 1.0 (mg/dl)						6		9	
WBC ≤ 4.5 ($\times 10^3/\text{mm}^3$)								10	
History of COPD							2	11	
Impaired sensorium			2	5	7			12	
Steroid use								13	
History of CHF		7							
Comp. score		8							
Race (white = 0)			9	4					
Type (lobec = 1)					1				
History of TIA				1					
CVA neurologic deficit				2					
Creatinine > 1.2 (mg/dl)			5	6					
ETOH			8						
PTT > 25 (sec)			7						
PT > 12 (sec)			4						
C-index	0.75	0.77	0.72	0.64	0.72	0.84	0.84	0.85	0.79

BUN, blood urea nitrogen; CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; CVA, cerebrovascular accident; DNR, do not resuscitate; ETOH, alcoholic; PTT, partial thromboplastin time; PT, prothrombin time; SGOT, aspartate transaminase; TIA, transient ischemic attack WBC, white blood cell count.

Table 5. PEARSON CORRELATION COEFFICIENT OF PROCEDURE AND SPECIALTY ANNUAL VOLUME

Procedure	Type of Volume	R	P
With risk-adjusted 30-day mortality			
Abdominal aortic aneurysmectomy	Procedure	-0.11	0.28
	Specialty	-0.10	0.29
Infrainguinal vascular reconstruction	Procedure	0.14	0.14
	Specialty	0.16	0.10
Carotid endarterectomy	Procedure	0.07	0.51
	Specialty	-0.02	0.84
Lobectomy/pneumonectomy	Procedure	-0.09	0.37
	Specialty	-0.06	0.53
Open cholecystectomy	Procedure	-0.14	0.13
	Specialty	-0.16	0.08
Laparoscopic cholecystectomy	Procedure	-0.03	0.73
	Specialty	-0.09	0.34
Colectomy	Procedure	-0.09	0.30
	Specialty	-0.05	0.57
Total hip arthroplasty	Procedure	0.09	0.35
	Specialty	0.01	0.89
With risk-adjusted 30-day stroke rate			
Carotid endarterectomy	Procedure	-0.13	0.18
	Specialty	-0.21	0.03

higher-volume quartiles. Although no interquartile differences in the O/E ratio were found after infrainguinal reconstruction, the expected and observed mortality rates after this procedure were significantly higher in the high-volume quartile hospitals compared to the low-quartile hospitals. This indicated that higher-risk patients underwent infrainguinal reconstruction at the higher-volume hospitals, underscoring the value of risk adjustment in the comparative assessment of outcome between hospitals with varying volumes.

Mixed Effects Hierarchical Modeling

In this modeling, 30-day mortality is the dependent variable and patient risk factors and procedure and surgical specialty volumes per hospital per year are the independent variables. After adjusting for the patient's risk in level 1 of the model, the procedure and the surgical specialty volumes per hospital per year were each entered into separate models to determine whether each was significantly associated with 30-day mortality in any of the eight procedures and 30-day stroke in CEA. The results of this modeling are shown in Table 7. The table lists for each operation the patient variables predictive of death (or stroke for one CEA analysis), arranged in the order of decreasing importance, along with the beta coefficient, the standard error, and the probability value for each variable. For each operation, the results of entering the two volume variables into the respective model are shown below the random intercept. In none of the nine models shown in the table was the procedure volume per hospital per year or the respective surgical specialty volume

per hospital per year a significant predictor of the risk-adjusted 30-day outcome.

To determine whether there was a threshold volume below which adverse outcomes were encountered in any of the eight procedures studied, the mixed effects hierarchical analysis was repeated with the quartiles of procedure and specialty volumes entered as independent variables. None of the models identified a specific volume quartile as a significant predictor of risk-adjusted 30-day outcomes, again confirming the lack of relation between hospital volume and outcome in these nine models.

Automatic Interaction Detection Analysis

When set for two ranges of volume, this analysis attempts to detect automatically a volume threshold that might significantly affect outcome. In all eight operations, no volume cutpoint was detected below which a significant increase in risk-adjusted 30-day mortality was observed. Likewise, in CEA, no volume cutpoint was detected below which a significant increase in the risk-adjusted 30-day stroke rate was observed.

DISCUSSION

Using a large prospective series of operations, we examined the relation between volume and risk-adjusted outcomes for eight operations commonly performed on inpatient surgical services in VA hospitals. The volume of each operation and the volume of surgery performed on the corresponding surgical specialty service were related to the risk-adjusted 30-day mortality rate in the eight operations and to the 30-day stroke rate in CEA. Patient risk factors and the models used for risk adjustment were developed from the NSQIP, which accrues prospectively patient risk factors and outcome information on patients undergoing major surgery on all surgical services in the VHA, using dedicated clinical nurse managers.¹⁹ Employing these models, the risk-adjusted 30-day mortality and morbidity rates have been shown to reflect the quality of surgical care in VA hospitals.²⁰

We used several analytic methods to evaluate the relation of volume to risk-adjusted outcome, including Pearson correlation coefficients between O/E ratios and volumes, analysis of variance comparing mean O/E ratios in groups defined by quartiles of procedure volume, hierarchical logistic regression, and automatic interaction detection. Mixed effects hierarchical logistic regression²⁹ takes into account both patient and hospital factors explanatory of death or stroke. In all of the analyses performed, we failed to document a relation between specialty and procedure volumes and risk-adjusted outcome. In one of the analyses (Pearson correlation), we observed a weak correlation between the specialty, not the procedure, volume and the risk-adjusted stroke rate after CEA. This relation, however, was not confirmed by the more appropriate hierarchical logistic regression analysis. We were also unable to dem-

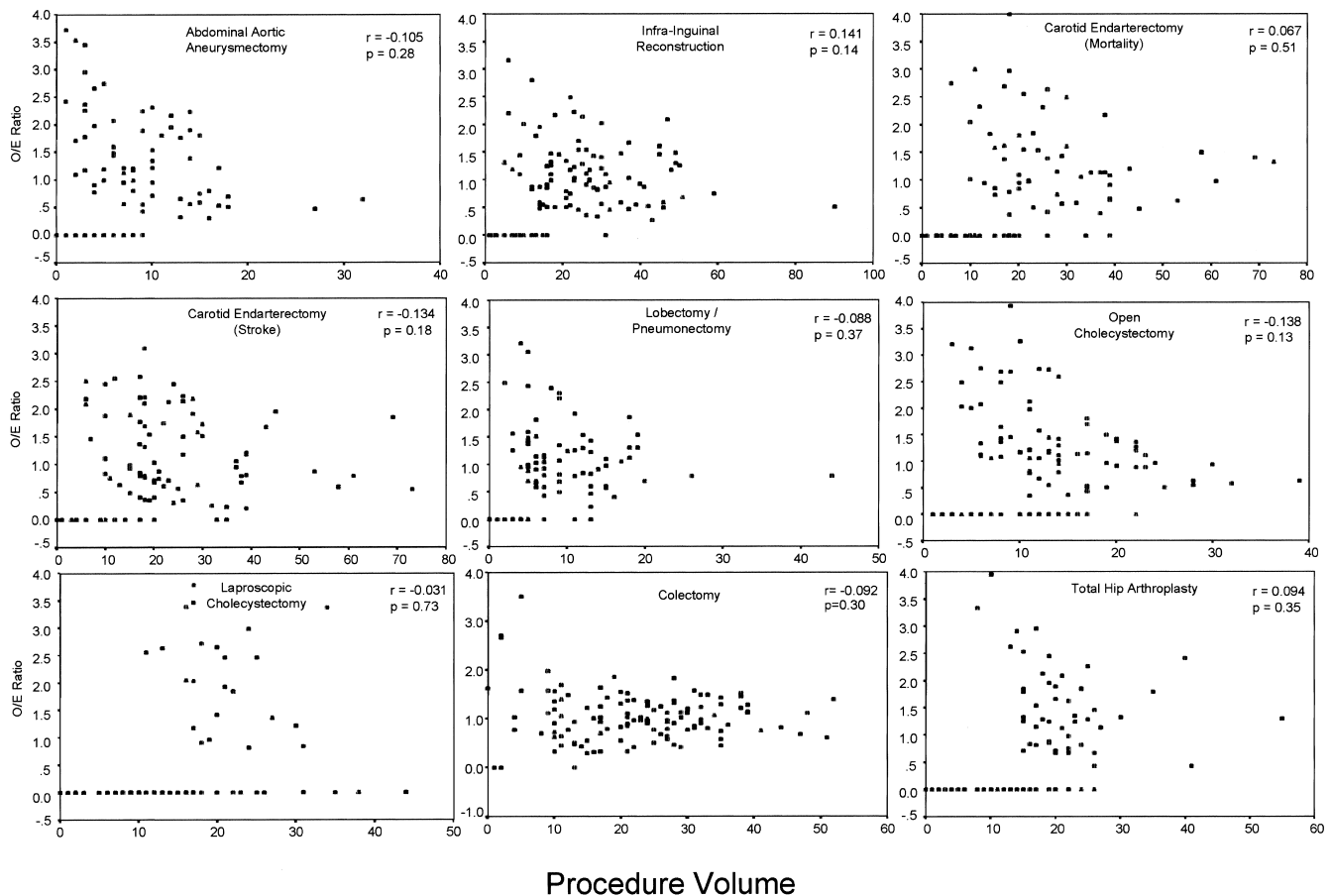


Figure 1. Relation of procedure volume to risk-adjusted 30-day mortality rate after eight operations and to risk-adjusted 30-day stroke rate after CEA. Each small circle represents a single VAMC. In each panel, the ordinate is the risk-adjusted outcome expressed as the O/E ratio, and the abscissa is the procedure volume expressed as the number of operations performed per year. The Pearson correlation coefficient (r) and the probability value for each relation are shown in the right upper corner of each panel. There were no significant correlations in any of the nine scattergrams shown on this figure.

onstrate in this study a threshold effect of volume below which poorer outcomes might be identified. In our analysis, the expected mortality rate was calculated based on the preoperative patient risk factors and, as such, was indicative of the severity of illness of the population of patients being analyzed. We found no difference in the severity of illness of patients undergoing surgery at high- and low-volume hospitals (except for infrainguinal vascular reconstruction, where higher-volume hospitals had sicker patients; see Table 6).

The relation between the volume of surgical operations and outcome has been debated for many years.³⁰ In recent years, analysis of this relation has focused on operations performed in the private sector that are technically complex and, in some cases, not commonly performed. Early reports argued for strong inverse relations between volume and outcome for coronary artery bypass surgery (CABG) and CEA.^{2,5,30,31} Similar relations have been suggested for pancreaticoduodenectomy,¹³⁻¹⁵ total hip arthroplasty,⁸ major vascular procedures including abdominal aortic aneurysm repair and CEA,^{3,4,7,9,11,32} and complex cancer resections.¹²

Most of these studies were retrospective and used state-sponsored registries (Table 8) in geographic areas where only one or a few high-volume centers existed. Because most of these studies were based on administrative databases with little clinical information (see Table 8), they were limited in their ability to risk-adjust adequately for the preoperative severity of illness.^{16,17}

Because of the limited risk adjustment and the skewing of the volume/outcome analysis by a few centers with very high volumes, two concerns are raised about previously published reports. First, it is not known whether high-volume centers in the private sector attract patients with levels of risk different from patients who seek care at lower-volume hospitals. Second, there has been no attempt to disentangle the influence of the expertise (*i.e.*, quality) from the experience (*i.e.*, volume) of the institutions or their individual surgeons. Unless these influences are recognized, there is a danger of attributing a lower operative mortality rate to high volume itself. Our study indicates that for prevalent operations in VA hospitals, sicker patients do not necessarily seek higher-volume institutions, and lower-vol-

Table 6. EXPECTED, OBSERVED, AND RISK-ADJUSTED RATES

Mortality Rates per Annual Quartiles of Procedure Volume				
Procedure	Quartiles of Procedure Volume/Hospital/Year	Expected Mortality (%)	Observed Mortality (%)	O/E Ratio
Abdominal aortic aneurysmectomy	Q1 = 0–3	4.7 ± 3.3	8.2 ± 17.3	1.75 ± 4.9
	Q2 = 4–6	5.3 ± 2.1	5.3 ± 5.6	0.92 ± 1.0
	Q3 = 7–10	4.8 ± 1.5	4.4 ± 3.0	0.93 ± 0.7
	Q4 = 11–32	4.3 ± 1.0	4.6 ± 2.7	1.08 ± 0.7
		P = 0.61	P = 0.49	P = 0.65
Infrainguinal vascular reconstruction	Q1 = 0–13	2.3 ± 0.7	1.9 ± 2.7	0.69 ± 1.0
	Q2 = 14–22	3.2 ± 0.8	3.0 ± 1.8	0.96 ± 0.6
	Q3 = 23–31	3.1 ± 0.6	3.4 ± 1.5	1.11 ± 0.5
	Q4 = 32–90	3.1 ± 0.5	3.0 ± 2.0	0.96 ± 0.5
		P = 0.001	P = 0.05	P = 0.15
Carotid endarterectomy	Q1 = 0–10	1.1 ± 0.4	0.8 ± 1.8	0.79 ± 2.1
	Q2 = 11–18	1.3 ± 0.4	1.4 ± 1.8	0.96 ± 1.2
	Q3 = 19–28	1.2 ± 0.2	1.2 ± 1.4	1.09 ± 1.2
	Q4 = 29–73	1.2 ± 0.2	1.2 ± 0.7	1.02 ± 0.6
		P = 0.14	P = 0.56	P = 0.89
Lobectomy/pneumectomy	Q1 = 0–5	5.1 ± 1.7	7.1 ± 19.1	1.28 ± 3.4
	Q2 = 6–7	5.5 ± 1.1	5.0 ± 4.9	0.90 ± 0.7
	Q3 = 8–13	5.3 ± 1.2	6.4 ± 3.2	1.27 ± 0.7
	Q4 = 14–44	5.5 ± 0.8	5.2 ± 3.0	0.91 ± 0.5
		P = 0.59	P = 0.85	P = 0.77
Open cholecystectomy	Q1 = 0–6	2.9 ± 1.6	3.8 ± 7.1	1.67 ± 5.3
	Q2 = 7–11	2.8 ± 1.1	4.1 ± 4.1	1.36 ± 1.2
	Q3 = 12–15	2.9 ± 1.5	2.7 ± 3.2	0.82 ± 0.9
	Q4 = 16–39	2.9 ± 1.3	2.6 ± 2.1	0.87 ± 0.5
		P = 1.0	P = 0.47	P = 0.61
Laparoscopic cholecystectomy	Q1 = 0–9	0.6 ± 0.6	0.6 ± 1.8	1.56 ± 5.5
	Q2 = 10–15	0.5 ± 0.2	0.5 ± 1.0	0.95 ± 2.4
	Q3 = 16–19	0.6 ± 0.3	0.6 ± 0.9	0.83 ± 1.3
	Q4 = 20–44	0.6 ± 0.3	0.6 ± 0.6	1.08 ± 1.4
		P = 0.60	P = 0.95	P = 0.82
Colectomy	Q1 = 0–12	7.9 ± 10.3	9.6 ± 17.4	1.26 ± 1.1
	Q2 = 13–22	6.9 ± 2.1	6.2 ± 3.2	0.90 ± 0.5
	Q3 = 23–30	6.7 ± 1.5	6.8 ± 2.9	1.00 ± 0.3
	Q4 = 31–52	7.2 ± 1.6	7.5 ± 2.9	1.05 ± 0.3
		P = 0.83	P = 0.46	P = 0.13
Total hip arthroplasty	Q1 = 0–10	1.7 ± 2.5	0.6 ± 2.0	0.80 ± 2.9
	Q2 = 11–16	1.2 ± 0.4	0.8 ± 1.2	0.66 ± 1.0
	Q3 = 17–22	1.1 ± 0.3	1.0 ± 0.8	0.98 ± 0.9
	Q4 = 23–55	1.0 ± 0.4	1.3 ± 0.9	1.43 ± 1.2
		P = 0.28	P = 0.40	P = 0.52

Stroke Rates per Quartiles of Procedure Volume/Hospital Year

Procedure	Quartiles of Procedure Volume/Hospital/Year	Expected Stroke Rate (%)	Observed Stroke Rate (%)	O/E Ratio
Carotid endarterectomy	Q1 = 0–10	1.9 ± 0.4	3.5 ± 6.0	1.72 ± 2.9
	Q2 = 11–18	2.1 ± 0.3	2.6 ± 1.8	1.20 ± 0.9
	Q3 = 19–28	2.1 ± 0.3	2.1 ± 1.5	1.03 ± 0.8
	Q4 = 29–73	2.1 ± 0.3	1.9 ± 1.4	0.88 ± 0.6
		P = 0.25	P = 0.30	P = 0.27

Table 7. TABLE OF MIXED EFFECTS HIERARCHICAL LOGISTIC REGRESSION MODELS

Procedure	Variables*	Estimate (β)	SE	Z	p Value
Aortic aneurysmectomy	Intercept	-6.37088	1.09852	-5.79949	0.00000
	Emergency	0.66231	0.32640	2.02912	0.04245
	ASA class	0.99793	0.14759	6.76171	0.00000
	WBC > 11.0 ($\times 10^3/\text{mm}^3$)	0.90487	0.23096	3.91781	0.00009
	BUN > 40 (mg/dl)	0.93544	0.36988	2.52905	0.01144
	Age (years)	0.03945	0.01121	3.51921	0.00043
	Weight loss	0.83552	0.41953	1.99155	0.04642
	Albumin (gm/dl)	-0.43927	0.17456	-2.51647	0.01185
	Random intercept	0.23315	0.33773	0.69034	0.24499
	Procedure volume	-0.02844	0.02133	-1.33340	0.10240
	Specialty volume**	-0.00196	0.00220	0.89150	0.37202
Infringuinal vascular reconstruction	Intercept	8.42033	0.82753	-10.17520	0.00000
	ASA class	0.84290	0.10993	7.66733	0.00000
	Emergency	0.76283	0.17126	4.45424	0.00001
	Age (years)	0.04803	0.00361	5.57802	0.00000
	Albumin (gm/dl)	-0.50747	0.10460	-4.85145	0.00000
	SGOT > 40 (IU/L)	0.71009	0.14184	5.00631	0.00000
	WBC > 11.0 ($\times 10^3/\text{mm}^3$)	0.45813	0.12804	3.57812	0.00035
	History of CHF	0.50986	0.18154	2.80848	0.00498
	Comp. Score	0.33686	0.08979	3.75150	0.00018
	Dyspnea	0.31341	0.11324	2.76764	0.00565
	BUN > 40 (mg/dl)	0.46238	0.23239	1.98970	0.04662
	Random intercept	0.29063	0.10289	2.82481	0.00237
	Procedure volume	-0.00181	0.00434	-0.41638	0.67713
	Specialty volume**	0.00011	0.00148	0.07197	0.94263
	Intercept	-6.26446	0.58456	-10.71657	0.00000
Carotid endarterectomy (mortality)	Emergency	1.48230	0.32214	4.60137	0.00000
	Impaired sensorium	1.22020	0.38070	3.20518	0.00135
	Functional status	0.61602	0.27306	2.25600	0.02407
	PT > 12 (sec)	0.66200	0.22143	2.98966	0.00279
	Creatinine > 1.2 (mg/dl)	0.56144	0.22225	2.52615	0.01153
	ASA class	0.54564	0.22240	2.45348	0.01415
	PTT > 25 (sec)	-0.52899	0.25002	-2.11573	0.03437
	ETOH	0.58243	0.30099	1.93507	0.05298
	Race (white = 0)	0.52922	0.26442	2.00144	0.04534
	Random intercept	0.20214	0.25891	0.78076	0.21747
	Procedure volume	0.00357	0.01000	0.35648	0.72148
	Specialty volume**	-0.00034	0.00247	-0.13920	0.88929
	Intercept	-4.49789	0.20684	21.74594	0.00000
	History of TIA	0.61076	0.13130	4.65180	0.00000
	CVA neuro deficit	0.45481	0.16431	2.76799	0.00564
Carotid endarterectomy (stroke)	Emergency	0.82933	0.36033	2.30155	0.02136
	Race (white = 0)	0.42947	0.21190	2.02677	0.04260
	Impaired sensorium	0.84300	0.40104	2.10203	0.03555
	Creatinine > 1.2 (mg/dl)	0.35824	0.15184	2.35935	0.01831
	Random intercept	0.34377	0.16903	2.03377	0.02099
	Procedure volume	-0.00338	0.00662	-0.51083	0.60947
	Specialty volume**	-0.00198	0.00189	-0.14587	0.29562
	Intercept	-3.63002	0.87998	-4.12513	0.00004
	Lobectomy	-1.10284	0.13651	-8.07895	0.00000
	Albumin (gm/dl)	-0.54265	0.12052	-4.50250	0.00001
	Age (years)	0.03909	0.00912	4.28751	0.00002
	ASA class	0.52507	0.12972	4.04775	0.00005
	SGOT > 40 (IU/L)	0.71418	0.23858	2.99338	0.00276
	BUN > 40 (mg/dl)	1.41801	0.49827	2.84587	0.00443
	Steroid use	0.66457	0.22363	2.97180	0.00296
Lobectomy/pneumonectomy	Dis. cancer	0.57592	0.29986	1.92061	0.05478
	Random intercept	0.00155	1.05082	0.00148	0.49941
	Procedure volume	-0.00866	0.01285	-0.67332	0.50075
	Specialty volume**	-0.00131	0.00356	-0.36661	0.71391

Table 7 (continued). TABLE OF MIXED EFFECTS HIERARCHICAL LOGISTIC REGRESSION MODELS

Procedure	Variables*	Estimate (β)	SE	Z	p Value
Open cholecystectomy	Intercept	-6.31811	0.98636	-6.40547	0.00000
	BUN > 40 (mg/dl)		0.23163	5.23733	0.00000
	Albumin (gm/dl)	-0.61702	0.12357	-4.99346	0.00000
	ASA class	0.78451	0.16728	4.68964	0.00000
	Age (years)	0.04193	0.00878	4.77403	0.00000
	Dyspnea	0.38489	0.13680	2.81351	0.00490
	Bilirubin > 1.0 (mg/dl)	0.38206	0.20994	1.81988	0.06878
	Functional status	0.35007	0.04601	2.39757	0.01650
	Weight loss	0.60208	0.33326	1.80664	0.07082
	SGOT > 40 (IU/L)	0.38774	0.18941	2.04706	0.04065
	Random intercept	0.22523	0.27396	0.82212	0.20550
	Procedure volume	-0.02828	0.01772	-1.59629	0.11042
	Specialty volume**	-0.00266	0.00149	-1.78578	0.07413
Laparoscopic cholecystectomy	Intercept	-1.88618	1.92949	-0.97755	0.32830
	Albumin (gm/dl)	-1.83760	0.28110	-6.53722	0.00000
	History of COPD	1.02587	0.38914	2.63627	0.00838
	Age (years)	0.04899	0.01851	2.64654	0.00813
	BUN > 40 (mg/dl)	1.52226	0.54619	2.78705	0.00532
	Functional status	0.57775	0.28981	1.99353	0.04620
	Random intercept	0.49569	0.35978	1.37775	0.08414
	Procedure volume	-0.02111	0.02359	-0.89482	0.37088
	Specialty volume**	-0.00314	0.00216	-1.45548	0.14554
	Intercept	-4.58706	0.45365	-10.11151	0.00000
Colectomy	ASA class	0.73754	0.06807	10.83549	0.00000
	Albumin (gm/dl)	-0.69684	0.06291	-11.07711	0.00000
	Emergency	0.75665	0.10992	6.88365	0.00000
	SGOT > 40 (IU/L)	0.64749	0.12625	5.12852	0.00000
	Age (years)	0.03365	0.00469	7.18232	0.00000
	Dis. cancer	0.67243	0.14195	4.73721	0.00000
	BUN > 40 (mg/dl)	0.41507	0.14032	2.95804	0.00310
	DNR	0.50506	0.19474	2.59347	0.00950
	Bilirubin > 1.0 (mg/dl)	0.30540	0.10190	2.99695	0.00273
	WBC ≤ 4.5 ($\times 10^6/\text{mm}^3$)	-0.45005	0.19350	-2.32588	0.02002
	History of COPD	0.26240	0.09746	2.69255	0.00709
	Impaired sensorium	0.40382	0.14644	2.75758	0.00582
	Steroid use	0.31238	0.15854	1.97039	0.04879
	Transfusion	0.25303	0.13803	1.83316	0.06678
	Random intercept	0.28971	0.07665	3.77946	0.00008
	Procedure volume	0.00157	0.00536	0.29229	0.77006
	Specialty volume	0.00084	0.00074	1.14377	0.25272
	Intercept	-5.47844	1.42012	-3.85774	0.00011
	Albumin (gm/dl)	-0.93138	0.23197	-4.01500	0.00006
	Dyspnea	0.73494	0.24428	3.00861	0.00262
Total hip arthroplasty	ASA class	0.79573	0.21518	3.69793	0.00022
	Age (years)	0.03660	0.01119	3.27163	0.00107
	SGOT > 40 (IU/L)	0.70600	0.44223	1.59646	0.11039
	Functional status	0.46225	0.22091	2.09246	0.03640
	Random intercept	0.14764	0.87760	0.25561	0.39913
	Procedure volume	0.01309	0.01106	1.18380	0.23649
	Specialty volume**	0.00054	0.00145	0.37301	0.70914
	Intercept				
	Albumin (gm/dl)				
	Dyspnea				

* In order of importance in the model.

** Two second level models were constructed. The first, "operation volume", includes the annualized volume of the operation for each hospital as the independent variable. The second, "specialty volume", includes the annualized volume of all major operations in that specialty as the independent variable.

BUN, blood urea nitrogen; CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; CVA, cerebrovascular accident; DNR, do not resuscitate; ETOH, alcoholic; PTT, partial thromboplastin time; PT, prothrombin time; SGOT, aspartate transaminase; TIA, transient ischemic attack; WBC, white blood cell count.

Table 8. CHARACTERISTICS OF STUDIES EVALUATING THE RELATIONSHIP BETWEEN OUTCOME AND HOSPITAL VOLUME IN EIGHT INTERMEDIATE COMPLEXITY OPERATIONS

Procedure	Author Year	Study Group	Time Interval	Endpoints	Risk Variable
AAA repair	Hannan 1989	NY State Reg. SPARCS	1986	In-hospital mortality	Age, gender; comorbid condition
AAA repair	Wen 1996	Ontario cooperative	1988–1992	In-hospital mortality	Age, gender; emergent status
AAA repair	Manheim 1998	California State Reg.	1982–1994	In-hospital mortality	Age; emergent status
Elective AAA repair	Sollano 1999	NY State Reg. SPARCS	1990–1995	30-day mortality	Age, gender; service intensity
Infrainguinal vasc. bypass	Manheim 1998	California State Reg.	1982–1994	In-hospital mortality	Age; emergent status
CEA	Perler 1998	Maryland State Reg.	1990–1995	In-hospital mortality, stroke	Age
CEA	Hannan 1998	NY State Reg.	1990–1995	In-hospital mortality	Age; comorbid condition
CEA	Wennberg 1998	Medicare hospitals in NASCET & ACAS trials	1992–1992	30-day mortality	Age; comorbid condition
CEA	Manheim 1998	California State Reg.	1982–1994	In-hospital mortality	Age; emergent status
Lobectomy/pneumonectomy	Romano 1992	California State Reg.	1983–1986	30-day mortality	Age, gender; comorbid condition
Pneumonectomy	Begg 1998	SEER/Medicare adm.	1984–1993	30-day mortality	Age; cancer stage; comorbid condition
Open Cholecystectomy	Hannan 1989	NY State Reg. SPARCS	1986	In-hospital mortality	Age, gender; comorbid condition
Colectomy	Hannan 1989	NY State Reg. SPARCS	1986	In-hospital mortality	Age; comorbid condition
Colectomy	Stremple 1993	HCUP-2/adm.	1984–1986	In-hospital mortality	Age; comorbid condition
THA	Taylor 1997	Medicare-administrative	1992–1994	30-day mortality	Age, gender

AAA, abdominal aortic aneurysmectomy; CEA, carotid endarterectomy; THA, total hip arthroplasty; Vasc., vascular.

ume institutions do not necessarily care for healthier patients. Most importantly, our findings do not support the hypothesis that with proper risk adjustment, being cared for in these higher-volume VAMCs for these prevalent intermediate-complexity procedures necessarily means better outcome. Hence, in VHA, and possibly in the private sector as well, the quality of surgical care is determined by hospital structure and processes,^{21,22} which do not include the volume of the surgery performed. Good risk-adjusted outcomes in high-volume hospitals reflect the high quality of surgical care, but high surgical volume alone is not necessarily the reason for the superior quality of surgical care in these hospitals.

What are the potential limitations of the current study? The first issue is whether death is an appropriate end point for evaluation of the relation between volume and outcome in these eight prevalent intermediate-complexity procedures. In prior reports from the NSQIP, we have documented that risk-adjusted 30-day mortality for all operations performed in a VA hospital was associated with the quality of surgical care at that hospital.^{21,22,24} This has prompted the VA to use the all-operations risk-adjusted 30-day mortality rate (the O/E ratio) as an ongoing comparative measure of the quality of surgical care in VHA.¹⁹ The relation of the specific operation O/E ratios reported in this study to the quality of surgical care in the various services within VHA has not yet been investigated. The c-indices generated by the mortality risk-adjustment models reported in this study (see Table 4) indicate good to excellent predictive validity of these risk-adjustment models and provide an advantage over other studies in the literature that have used unadjusted or poorly adjusted operative mortality rates to compare the relation of surgical volume to outcome (see Table 8).

Except for postoperative stroke after CEA, this study has not attempted to explore the relation between surgical volume and postoperative risk-adjusted adverse occurrences, which the NSQIP has also shown to be associated with the quality of surgical care.²⁰ The NSQIP is currently investigating this relation, which will be the subject of a separate publication. Recent studies have demonstrated an inverse relation between surgical volume and the incidence of postoperative adverse occurrences^{33,34} and a direct relation between surgical volume and long-term cancer-free survival.³⁵

The second potential limitation of this study is whether operative volume in VAMCs is comparable to that reported from the private sector. A direct comparison between volumes of these operations in VHA and the private sector cannot be easily made or interpreted. Information from private sector registries or trials is limited because surgical volumes are not often expressed in relation to the institution, but in relation to individual surgeons. Information in the NSQIP database is currently limited to institutional volumes. In addition, many VA surgeons do not limit their practices to VA hospitals. Hence, VA hospital volume does not necessarily correlate with the entire experience and

surgical volume of the VA surgical staff. With these limitations, the data in Table 3 indicate that procedure volumes in VA hospitals generally correlate with the ranges published for private sector hospitals with low and intermediate volumes. With the possible exception of abdominal aortic aneurysm repair⁴ and infrainguinal reconstruction, few of the VA hospitals have procedure volumes matching high-volume institutions in the private sector.

A third potential limitation of this study is that a reduced set of preoperative patient risk factors was collected at the low-volume VAMCs as a result of the lack of funding for a surgical clinical nurse reviewer. These VAMCs provided the data to the NSQIP through use of their own resources. Although this reduced data set did not affect the prediction models for death for seven of the eight operations, it did affect the prediction models for death and stroke for CEA. Some of the important predictor variables for this operation were not included in the reduced data set. Therefore, in the analysis involving CEA, 8 low-volume VAMCs were excluded from the total of 101 VAMCs in which this operation was performed.

Despite the absence of a large number of high-volume hospitals in VHA, the unadjusted mortality rates in our study are generally comparable to those that have been reported in state-wide databases, among Medicare patients, or among patients treated in medical centers that participate in clinical trials. For total hip arthroplasty, we report a mortality rate of 1% (see Table 2); the HCUP-2 database from the mid-1980s reported a mortality rate of 1%,³⁶ as did the HCFA database derived from Medicare patients in the early 1990s for the highest-volume institutions.⁸ For elective or semielective aortic aneurysm repair, the VHA overall mortality rate of 4.7% can be compared with mortality rates of 3.8% to 6% reported from various registries in California, New York State, Ontario, and Scandinavia.^{3,6,7,37} For infrainguinal vascular bypass procedures, the VHA overall mortality rate of 3.1% can be compared with a mortality rate of 3.3% observed in the California registry.⁷ For lobectomy/pneumonectomy, there are no clear differences compared with data reported from administrative registries.^{5,12} Outcomes for CEA, both in terms of mortality rate (1.2%) and stroke rate (2.1%), in VHA are comparable to those reported for intermediate- to high-volume institutions recorded in the state registries of Connecticut, California, Maryland, and New York.^{7,9-11,32} For open cholecystectomy, VHA mortality rates are within the ranges reported from prior HCFA data.^{5,36}

Comparable, population-based figures are not available for laparoscopic cholecystectomy because reports from the private sector³⁸ reflect a fundamentally different patient population. In a detailed analysis of VHA outcomes for open and laparoscopic cholecystectomy,³⁹ we observed that several key factors contributed to mortality for these specific procedures, including emergency surgery, activities of daily living indices, serum albumin level, and ASA risk. We predict that when adjusted for risk, outcomes of laparo-

scopic cholecystectomy in VHA patients would be similar to those that would be observed in the private sector, as previously suggested.³⁶

Despite concerns regarding differences in surgical patient characteristics between VHA and the private sector, the above published data indicate that both volume and outcomes for the eight prevalent intermediate-complexity procedures reported here for VA hospitals are comparable to those in the private sector. The results of this study may not be applicable to the ongoing debate about whether procedures of high complexity should be performed at regionally designated referral centers. Perhaps the two most publicized examples of such procedures are CABG and pancreaticoduodenectomy.

In CABG, it is clear that a practice, rather than an individual surgeon, needs to perform approximately 100 CABG procedures annually to have early postoperative results equivalent to centers with much higher volume.^{4,40,41} A recent study based on the 1990 to 1995 New York State cardiac surgery database failed to show a relation between volume of CABG procedures performed and the risk-adjusted mortality rate.⁴ It confirmed an earlier report based on an analysis of nearly 24,000 CABG procedures performed between 1987 and 1992 at 44 VAMCs, all of which are currently participating in the NSQIP. This study also showed no relation between annual hospital CABG procedure volume and the risk-adjusted mortality rate.⁴¹

In pancreaticoduodenectomy, studies using two statewide registries demonstrated improved outcome with higher volume, leading to pleas for regionalization so that only certain high-volume centers would perform this procedure in the private sector.^{13–15} The results of our study do not bear directly on these highly technical and complex procedures, except to provide a warning that fair and meaningful assessment of the relation of outcomes and procedure volume will require appropriate risk adjustment.

The findings of this study are based on a specific patient population—older, economically disadvantaged, predominantly male veterans—who may not be comparable to other patient populations. Our findings may have significant implications for other large integrated health care delivery systems with regard to the deployment of and accessibility to surgical care. If there are no volume/outcome-related differences for simple or intermediate-complexity surgical cases, then there may be no clinical reason to curtail the availability of this type of surgical care at smaller or lower-volume hospitals. From access to care and patient satisfaction perspectives, this would be highly desirable.

Our findings, however, need to be validated by others before VHA results can be generalized to other settings. The demographics and generally higher risk status (*i.e.*, adverse selection) of VA patients and the significant role that residents assume in the care of VA patients may accentuate any direct volume/outcome-related relation. Conversely, the high degree of academic affiliation of VA hospitals and the ongoing non-VA surgical experience of VA surgeons may

obscure such differences. Despite the intriguing nature of our findings, generalization of the VHA experience to other integrated health care delivery systems must be tempered until the findings are validated by others studying such large health care delivery systems.

In summary, we have shown that in VA hospitals, the procedure and surgical specialty volumes in eight prevalent operations of intermediate complexity are not associated with the risk-adjusted 30-day mortality rate from these operations, or with the risk-adjusted 30-day stroke rate after CEA. Until convincing evidence is provided to the contrary (using risk-adjusted 30-day morbidity rates or other validated quality indicators), surgical volume should not be used as a surrogate measure of quality in an integrated health care delivery system such as VHA.

Acknowledgments

The authors thank Nancy Healey, Cindy Bofetiado, Craig Miller, Laki Rousou, and Gail Meyberger for their help in preparing this manuscript.

References

1. Kizer KW. The "New VA": a national laboratory for health care quality management. *Am J Med Quality* 1999; 14(1):3–20.
2. Hannan EL, Kilburn H, Bernard H. Coronary artery bypass surgery: the relationship between in-hospital mortality rate and surgical volume after controlling for clinical risk factors. *Med Care* 1991; 29:1094–1097.
3. Hannan EL, Kilburn H Jr, O'Donnell JF, et al. A longitudinal analysis of the relationship between in-hospital mortality and the volume of abdominal aortic aneurysm surgeries performed. *Health Serv Res* 1992; 27:517–542.
4. Sollano JA, Gelijns AC, Moskowitz AJ. Volume-outcome relationships in cardiovascular operations: New York State 1990–1995. *J Thorac Cardiovasc Surg* 1999; 117:419–430.
5. Hannan EL, O'Donnell JF, Kilburn H Jr, et al. Investigation of the relationship between volume and mortality for surgical procedures in New York State hospitals. *JAMA* 1989; 262:503–510.
6. Wen SW, Simunovic M, Williams JI, Johnston KW, Naylor CD. Hospital volume, calendar age, and short-term outcomes in patients undergoing repair of aortic aneurysms: the Ontario experience, 1988–92. *J Epidemiol Community Health* 1996; 50:207–213.
7. Manheim LM, Sohn MW, Feinglass J, et al. Hospital vascular surgery volume and procedure mortality rates in California. *J Vasc Surg* 1998; 28:45–56.
8. Taylor HD, Dennis DA, Crane HS. Relationship between mortality rates and hospital volume for Medicare patients undergoing major orthopedic surgery of the hip, knee, spine and femur. *J Arthroplasty* 1997; 12:235–242.
9. Perler BA, Dardik A, Burleyson GP, Gordon TA, Williams GM. Influence of age and hospital volume on results of carotid endarterectomy: a state-wide analysis of 9918 cases. *J Vasc Surg* 1998; 27:25–31.
10. Hannan EI, Popp AJ, Tranmer B, Fuestel P, Waldman J, Shah D. Relationship between provider outcome and mortality for carotid endarterectomies in New York State. *Stroke* 1998; 29:2292–2297.
11. Wennberg DE, Lucas FL, Birkmeyer JD, Bredenberg CE, Fisher ES. Variation in carotid endarterectomy mortality in the Medicare population: trial hospitals, volume, and patient characteristics. *JAMA* 1998; 279:1278–1281.

12. Begg CB, Cramer LD, Hoskins WJ, Brennan MF. Impact of hospital volume on operative mortality for major cancer surgery. *JAMA* 1998; 280:1747-1751.
13. Lieberman MD, Kilburn H, Lindsey M, Brennan MF. Relation of perioperative deaths to hospital volume among patients undergoing pancreatic resection for malignancy. *Ann Surg* 1995; 222:638-645.
14. Birkmeyer JD, Finlayson SRG, Tosteson ANA, Sharp SM, Warshaw AL, Fisher ES. Effect of hospital volume on in-hospital mortality with pancreaticoduodenectomy. *Surgery* 1999; 25:250-256.
15. Sosa JA, Bowman HM, Gordon TA, et al. Importance of hospital volume in the overall management of pancreatic cancer. *Ann Surg* 1998; 228:429-438.
16. Iezzoni LI, Daley J, Heeren T, et al. Identifying complications of care using administrative data. *Medical Care* 1994; 32:700-715.
17. Iezzoni LI. Data sources and implications: administrative databases. In: Iezzoni LI, ed. *Risk adjustment for measuring healthcare outcomes*. Ann Arbor: Health Administration Press; 1997:169-242.
18. Khuri SF, Daley J, Henderson WG, et al. The National VA Surgical Risk Study. Risk adjustment for the comparative assessment of the quality of surgical care. *J Am Coll Surg* 1995; 180(5):519-531.
19. Khuri SF, Daley J, Henderson WG, et al. The Department of Veteran Affairs' NSQIP: the first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. *Ann Surg* 1998; 228:491-507.
20. Daley J, Khuri SF, Henderson WG, et al. Risk adjustment of the post-operative morbidity rate for the comparative assessment of the quality of surgical care. *J Am Coll Surg* 1997; 185:327-340.
21. Young GJ, Charns MP, Daley J, Forbes M, Henderson W, Khuri SF. Best practices for managing surgical services: the role of coordination. *J Healthcare Management Rev* 1997; 22(4):85-99.
22. Young GJ, Charns MP, Desai K, et al. Patterns of coordination and clinical outcomes: a study of surgical services. *Health Serv Res* 1998; 33:1211-1236.
23. Khuri SF, Daley J, Henderson WG, et al. Risk adjustment of the post-operative mortality rate for the comparative assessment of the quality of surgical care. *J Am Coll Surg* 1997; 185:315-327.
24. Daley J, Forbes MG, Young GJ, et al. Validating risk-adjusted surgical outcomes: site visit assessment of process and structure. *J Am Coll Surg* 1997; 185:341-351.
25. Morgan JA, Sonquist JA. Problems in the analysis of survey data, and a proposal. *J Am Statistical Assoc* 1963; 58:415-434.
26. Sonquist JA, Morgan JN. The detection of interaction effects: a report on a computer program for the selection of optimal combinations of explanatory variables. Ann Arbor: Institute for Social Research, University of Michigan; 1964.
27. PC Group Users Guide, Version 3.01, 1992. Austin Data Management Associates.
28. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic curve. *Radiology* 1982; 143:29-36.
29. Wong GY, Mason WM. The hierarchical logistic regression model for multilevel analysis. *J Am Stat Assoc* 1985; 80:513-524.
30. Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empiric relation between surgical volume and mortality. *N Engl J Med* 1979; 301:1364-1369.
31. Sowden AJ, Deeks JJ, Sheldon TA. Volume and outcome in coronary artery bypass graft surgery: true association or artifact? *Br Med J* 1995; 311:151-155.
32. Ruby ST, Robinson D, Lynch JT, Mark H. Outcome analysis of carotid endarterectomy in Connecticut: the impact of volume and specialty. *Ann Vasc Surg* 1996; 10:22-26.
33. Sosa JA, Bowman HM, Tielsch JM, Powe NR, Gordon TA, Udelsman R. The importance of surgeon experience for clinical and economic outcomes from thyroidectomy. *Ann Surg* 1998; 228:320-330.
34. Sosa JA, Powe NR, Levine MA, Udelsman R, Zeiger MA. Profile of a clinical practice: thresholds for surgery and surgical outcomes for patients with primary hyperparathyroidism: a national survey of endocrine surgeons. *J Clin Endocr Metab* 1998; 83:2658-2665.
35. Porter GA, Soskolne CL, Yakimets WW, Newman SC. Surgeon-related factors and outcome in rectal cancer. *Ann Surg* 1998; 227:157-167.
36. Stremple JF, Bross DS, Davis CL, McDonald GO. Comparison of post-operative mortality in VA and private hospitals. *Ann Surg* 1993; 217:277-285.
37. Amundsen S, Skjaerven R, Trippestad A, Soreide O. Abdominal aortic aneurysms. Is there an association between surgical volume, surgical experience, hospital type and operative mortality? *Acta Chir Scand* 1990; 156:323-327.
38. Shea JA, Healey MJ, Berlin JA, et al. Mortality and complications associated with laparoscopic cholecystectomy. *Ann Surg* 1996; 224: 609-620.
39. Chen AY, Daley J, Pappas TN, Henderson WG, Khuri SF. Growing use of laparoscopic cholecystectomy in the National Veterans Affairs Surgical Risk Study. *Ann Surg* 1998; 227(1):12-24.
40. Clark RE. Outcome as a function of annual coronary artery bypass graft volume. The Ad Hoc Committee on Cardiac Surgery Credentialing of the Society of Thoracic Surgeons. *Ann Thorac Surg* 1996; 61:21-26.
41. Shroyer ALW, Marshall G, Warner BA, et al. No continuous relationship between Veterans Affairs hospital coronary artery bypass grafting surgical volume and operative mortality. *Ann Thorac Surg* 1996; 61:17-20.

Appendix

Current Program Participants

Chairperson's Office: Shukri F. Khuri, MD (Chairperson); Jennifer Daley, MD (Co-Chairperson); Jeannette Spencer, RN, MS, CS (National Clinical Coordinator); Nancy Healey, BS (Research Health Scientist); Craig Miller, BSEE (Clinical Information Specialist); Gail Mayberger (Secretary).

Hines Cooperative Studies Program Coordinating Center: William Henderson, PhD (Director); James Gibbs, PhD (Health Services Researcher); Kwan Hur, MS (Biostatistician); Bharat Thakkar (Statistical Programmer); Robbin Denwood, RN MSN, MBA (Data Coordinator).

Center for Continuous Improvement in Cardiac Surgery: Karl Hammermeister, MD (Chairman); Frederick Grover, MD (Co-Chairman); Randy Johnson, MS (Computer Specialist); Laurie Shroyer, PhD (Health Research Scientist).

Birmingham Information Systems Center: Alan Monosky (Computer Specialist); Michael Montali (Senior Systems Analyst); Steve Musgrove (Computer Specialist).

Executive Committee: Shukri F. Khuri, MD (Chairperson), Brockton/West Roxbury, MA; Jennifer Daley, MD (Co-Chairperson), Brockton/West Roxbury, MA; J. Bradley Aust, MD, San Antonio, TX; Richard H. Bell, Jr., MD, Seattle WA; John Demakis, MD, Washington DC; Vernon Chong, MD, Grand Prairie, TX; Peter J. Fabri, MD, Tampa, FL; James Gibbs, PhD, Hines, IL; Frederick Grover, MD, Denver, CO; Karl Hammermeister, MD, Denver, CO; William Henderson, PhD, Hines, IL; Gerald McDonald, MD, Washington DC; Edward Passaro, Jr., MD, West Los Angeles, CA; Lloyd Philips, MD, PhD, Ann Arbor, MI; Frank Scamman, Iowa City, IA; Jeannette Spencer, MS, RN, Brockton/West Roxbury, MA; John Stremple, MD, Pittsburgh, PA.

VA Medical Centers, Chiefs of Surgery, and Surgical Clinical Nurse Reviewers: Albany, NY: Thomas K. Wu, MD; George A. Leamy, RN; Albuquerque, NM: Stuart Ford, MD; Tony L. Lantzer, BSN; Alexandria, LA: Hollis Reed, MD; Carol Rowe, BSN; Altoona, PA: Akbar M. Samii, MD; Phyllis Podrasky, RN; Amarillo, TX: Mario Feola, MD; Sheree Keil, RN; Ann Arbor, MI: John F. Sweeney, MD; Linda S. Brooks, RN; Asheville, NC: Thomas J. Berger, MD; Marge Turcot, RN; Atlanta (De-

catur), GA: Aaron S. Fink, MD; Renee Lawrence, RN; Augusta, GA: George I. Cue, MD, Connie Miller, RN; Baltimore, MD: Barbara Bass, MD; Nancy P. Specht, RN, BSN, MAS; Bay Pines, FL: Terry Wright, MD; Judith M. Girard, RN; Beckley, WV: Georges A. Hoche, MD; Pam Johnson; Big Spring, TX: Gaddum Reddy, MD; Jennan Swafford, RNC, MSN; Biloxi, MS: Larry Fontenelle, MD; Donna Wells, RN; Birmingham, AL: John J. Gleysteen, MD; Linda Helm-Little, RN; Boise, ID: Ernest C. Peterson, MD; Launa J. Nardella, BSN; Boston, MA: Willard Johnson, MD; Laura McDonald, RN, BSN; Bronx, NY: A. James McElhinney, MD; Elias Enriquez, RN; Brooklyn, NY: Bimal C. Ghosh, MD, Wendy R. Trimboli, BSN, MA, CPHQ; Buffalo, NY: Irineo Gutierrez, MD; Mary Ann Blake, RN, MSN; Castle Point, NY: A. James McElhinney, MD; Barbara Powers, RN; Charleston, SC: John Allison, MD; Stephen E. Johnston, RN, MSN, CNA; Cheyenne, WY: D. Michael Kilpatrick, MD; Nina J. Pike, RN, BSN; Chicago (Lakeside), IL: Robert V. Rege, MD; Denise Ostrowski, RN; Chicago (Westside), IL: Donald K. Wood, MD; Carben Daniels, RN; Cincinnati, OH: Robert A. Bower, MD; Elaine Hardin, RN; Clarksburg, WV: Juanito V. Chua, MD; Lisa R. Michael, RN; Cleveland, OH: John Raaf, MD; Mary Ann Bobulsky, RN, BSN; Columbia, MO: Debra Koivunen, MD; Barbara Von Thun, RN, MSN; Columbia, SC: John Jeffrey Brown, MD; Joanne K. Ogg, RN, MN, CS; Dallas, TX: Richard H. Turnage, MD; Bernice Willis, RN; Danville, IL: Jin Kim, MD; David Lohnes, RN, BSN; Dayton, OH: Samuel A. Aderonojo, MD; Shirley Ribak, RN, MSN; Denver, CO: Frederick L. Grover, MD; Donna LoSaso, RN; Des Moines, IA: David Sidney, MD; Cathy S. Sandle, RN, BSN; Detroit, MI: Robert Kozol, MD; Barbara L. Bieke, RN, MSN; Dublin, GA: Noel Nellis, MD; Teresa Fagan, RN; Durham, NC: Theodore Pappas, MD; C. Jean Hanchey, RN; East Orange, NJ: Frank Padburg, MD; Anna T. Detschel, RN; Erie, PA: Prabhu Negi, MD; Denise Albertson, RN; Fargo, ND: William Kirby, MD; Priscilla K. Stroh, RN; Fayetteville, AR: Pat O'Donnell, MD; Carol Wolgamott, RN; Fayetteville, NC: Arthur McGuire, MD; Nancy Albaladejo, RN, MSA, CNA; Fort Harrison, MT: Michael Agee, MD; Edna L. Clausen, BSN; Fort Meade, SD: Mark F. Blum, MD; Teresa Gabeline, RN; Fort Wayne, IN: Sun Guo, MD; Fresno, CA: Gregory Wille, MD; Elena J. Eaton, RN, BSN; Gainesville, FL: Timothy Flynn, MD; Linda D. Carter, RN; Grand Island, NE: Danitsu Hirar, MD; Cynthia E. Hansen, RN; Grand Junction, CO: Earl Howells, MD; Karen L. Rogers, RN, BA; Hampton, VA: Ali Farpour, MD; Sol F. Aquinaldo, RN, BSN; Hines, IL: Charles H. Andrus, MD, FACS; Kristine L. Johnson, RN, BSN, BA, BS; Hot Springs, SD: Mark F. Blum, MD; Marcia Bishop, RRA; Houston, TX: Kamal Itani, MD; Barbara J. Anderson, LPN; Clara Kistner, RN; Huntington, WV: Timothy Canterbury, MD; Rena K. Black, RN; Indianapolis, IN: Dolores Cikirt, MD; Connie Adams, RN; Iowa City, IA: Kimberly S. Ephgrave, MD; Isabelle A. Olson, RN; Iron Mountain, MI: Robert L. Alexander, MD; Terri M. Danielson, RN, AND; Jackson, MS: Kenneth Simon, MD; Sheila Ann Buck, RN, BSN; Kansas City, MO: Mary McAnew, MD, and Betty Drees, MD; Becky Ganaban, RN, Berta Graves, RN; Kerrville, TX: Mauro Gangai, MD; Theresa Rangel, BS, MA, EdPsy; Lake City, FL: Giora Katz, MD; Charlotte L. Lintz, RN, MSN; Leavenworth, KS: Chris C. Haller, MD; Mary Lee Driscoll, RN, BSN, CNOR; Lebanon, PA: George Miash, MD; Vicki L. Leibich, LPN; Lexington, KY: Thomas Schwarcz, MD; Rose Mary Collins, RN; Lincoln, NE: Danitsu Hirar, MD; Judy L. Sanne, RN; Little Rock, AR: Nicholas P. Lang, MD; Richard D. Bloesch, RN; Loma Linda, CA: Sherif Emil, MD; Gillian Gomulka, RN; Long Beach, CA: Edward A. Stemmer, MD; Delores Whalen, RN; Louisville, KY: Richard Neal Garrison, MD; Ruth A. Meadows, RN; Madison, WI: James Starling, MD; Kathy Gruber, RN; Manchester, NH: Willard Johnson, MD; Patricia M. Stevens, RN, BSN; Marion, IL: Rama Iyengar, MD; Jane A. Hale, RN; Martinez, CA: Pauline Velez, MD; Marytess Baula, RN; Martinsburg, WV: C. R. Kamath, MD; Jeffrey B. Spoon, PA, BS; Memphis, TN: Eugene Mangiante, MD; Anita L. Garrison, RN; Miami, FL: A. J. Furst, MD; Nancy Box, RN; Milwaukee, WI: Charles Abrahamian, MD; Christine A. Tyler, RN, BSN; Minneapolis, MN: Donald G. McQuarrie, MD; Jane Bonawitz-Conlin, RN; Montgomery, AL: Eddie Warren, MD; Jeulia E. Hendrick, RN, BSN, MS; Mountain Home, TN: David Walters, MD; Joyce F. Hamm, RN, BSN; Murfreesboro, TN: Rudolph Comberbatch, MD, FACS, Margaret Cantrell, RN; Musko-

gee, OK: Dala Jarolin, MD; Terry Maycher, RN; Nashville, TN: Walter H. Merrill, MD; Jeanette B. Pujol, BSN, RNC; New Orleans, LA: Paul R. Hastings, MD, FACS; Donna M. Gray, RNC, CCRN; New York, NY: Alex C. Solowey, MD; Jacqueline H. Parker, RN, BSN; North Chicago, IL: B. F. Kepley, MD; Kathleen Mega, BSN; Northport, NY: Eugene P. Mohan, MD; Sheila Dahl, RN, MSN; Oklahoma City, OK: Donald R. Carter, MD; Rouchelle Osborn, RN, MS; Omaha, NE: Thomas Lynch, MD; Sharon M. Tighe, RN; Palo Alto, CA: Thomas Burdon, MD; Jacie Epperson, RN, BSN; Philadelphia, PA: Steven Raper, MD, Miriam S. Moskowitz, RN, MSN; Phoenix, AZ: William Doler, MD, FACS; Seaton West, RN; Pittsburgh, PA: John Stremple, MD; Susan Layne, RN; Portland, OR: Cliff Deveney, MD; Elizabeth McCollum, RN; Providence, RI: Michael P. Vezeridis, MD; Carol Maynard, RN; Reno, NV: Ralph DePalma, MD; Jerylann E. Gale; Richmond, VA: Hunter Holmes McGuire, Jr., MD; Gail Laub, RN; Roseburg, OR: Norman Marshall, MD; Dennis Morehouse, RN; Saginaw, MI: Isa Salti, MD; St. Louis, MO: Frank E. Johnson, MD; Mary Louise Smith, RN, MSN; Salem, VA: Wayne H. Wilson, MD; Rebecca P. Evans, RN, BSN; Salisbury, NC: Barbara Temmeck, MD; Lisa H. Noonan, RN, BSN, CPHQ; Salt Lake City, UT: Leigh Neumayer, MD; Sandy McMaster, RN; San Antonio, TX: O. LaWayne Miller, MD; Linda M. Porazzi, RN, MSN, CNS; San Diego, CA: Nicholas A. Halasz, MD; Gail P. Maxwell, RN, BA; San Francisco, CA: Jeffrey Norton, MD; Rita J. Sears, RN, MS; San Juan, PR: Ernesto Rive-Mora, MD; Saribelle Reyes-Frau, RN, BSN; Seattle, WA: Richard Bell, MD; Julie K. Kieras, RN, BSN; Sepulveda, CA: Howard Reber, MD; Betty Wright, RN; Shreveport, LA: James Evans, MD; Lillian Thornhill, RN; Sioux Falls, SD: John Ryan, MD; Becky Poss, RN; Spokane, WA: A. Garth Davis, MD; Shelly Sumner, ART, CPIM; Syracuse, NY: Michael Sobel, MD, John E. LeBeau, RN; Tacoma, WA: Richard Bell, MD; Julie K. Kieras, RN, BSN; Tampa, FL: Peter J. Fabri, MD; Kathryn S. Bowns, RN; Temple, TX: P. Pandya, MD; Carolyn Broussard, RN; Togus, ME: Martyn Vickers, MD; Patricia A. Wotton, RN; Topeka, KS: C. N. Radhakrishna, MD; Allen Zander, RN; Tucson, AZ: Martin L. Dresner, MD; Christopher R. Brown, RN, BSN; Tuskegee, AL: Eddie Warren, MD; Joice Promisee, RN, BSN; Washington DC: Anton Sidawy, MD; Deborah T. Fleming, RN, BSN; West Haven, CT: Ronnie Rosenthal, MD; Kathy Maher-Cleary, RN; West Los Angeles, CA: Edward Livingston, MD; Marilyn DeGroot, RN; West Palm Beach, FL: James Schell, MD; Rosa Caraballo, RN; West Roxbury, MA: Shukri Khuri, MD; Jeannette Spencer, RN, MS, CS; White River Junction, VT: Martha McDaniel, MD; Lisa Ryder, RN; Wichita, KS: Joseph K. Robertson, MD; Stephanie Lentz, RN; Wilkes-Barre, PA: Feroz Sheikh, MD; Beth A. Chaken, RN, MSN; Wilmington, DE: Claude Lieber, MD; Evie Logue, RN.

Discussion

DR. MURRAY F. BRENNAN (New York, New York): I would like to thank Dr. Khuri for the invitation to comment on the paper and appreciate receiving the manuscript ahead of time. Dr. Khuri is to be congratulated on his leadership role in the VA National Surgical Quality Improvement Program.

The present study of eight major surgical procedures, totaling 68,000 operations, relates procedure volume and its volume per hospital to mortality but does not show a correlation. This is very surprising to me, that the risk for all eight procedures of 30-day mortality was not dependent on volume, something that has been shown in the past by many others, including the previous paper by Dr. Harmon. (Slide.) This is a further analysis of over 53,000 operative cancer cases in New York City in 1 year. Both surgeon volume and hospital volume are compounding factors in overall mortality. I would emphasize that these data are risk-adjusted, and they suggest, in contradiction to Dr. Harmon and Dr. Khuri, that for cancer cases, hospital and surgical volume are additive for all

levels. An increase in hospital or in surgeon volume results in a decrease in mortality.

My question, therefore, is whether Dr. Khuri has any further explanation for his findings, which are at variance from other published studies? Is it simply that they have taken operations which in the main are common to all general and vascular surgeons and not the complicated procedures referred to in the past? This would seem to be unlikely given the data I have shown, which includes all cancer cases, and the data that Dr. Harmon has shown, which includes colectomy. Is it that the VA is a group with uniformly high standards, perhaps due to the quality improvement program initiated by Dr. Khuri? Or does it mean that the VA hospitals are associated with medical schools that have high-volume faculty?

Finally, Dr. Khuri answered the question of how many are too few. And I heard his answer, "there is no number that is too few." Do you believe that, Dr. Khuri? Would you be the first case by the first surgeon who does an occasional case?

PRESENTER DR. SHUKRI F. KHURI (West Roxbury, Massachusetts): I think there are a number of differences between the State of New York registry and the NSQIP database, two of which are relevant to your questions, Dr. Brennan.

The first relates to the degree of complexity of the operations which you are comparing in the two databases. Our study specifically targeted operations of intermediate complexity which were frequently performed in the VA. The study was conducted, in part, to determine whether the VA, as it faced decreasing budgets, would be justified in closing surgical services whose surgical volumes were relatively low. The data in this study clearly caution against using volume as the indicator for whether or not a certain surgical service should be closed. The complexity and type of these commonly performed operations in our study might not be the same as the complexity and type of the cancer operations in the registry in the State of New York. In fact, a recent publication in the *Journal of Thoracic and Cardiovascular Surgery* which was based on the New York State database showed no relationship between surgical volume and outcome in coronary artery bypass surgery—an operation which is highly complex.

The second difference between the New York State database and the NSQIP database is the predictability of the risk adjustment models. The NSQIP database is a clinical database collected prospectively by dedicated nurses. As I showed you today, the risk adjustment models for the various operations studied had high c-indices, indicating good predictability. Administrative databases, on the other hand, are mostly based on charge codes and might not have enough clinical information to provide highly predictable risk adjustment models—a fact which Dr. William Best, from our group, has recently demonstrated by comparing risk factors and outcomes contained in the NSQIP database to those contained in the VA Patient Treatment File.

DR. TOBY A. GORDON (Baltimore, Maryland): I would like to compliment Dr. Khuri and his colleagues for their efforts in leading and conducting outcomes research studies in the Veterans Administration. This paper in particular benefits from the large sample size, the large number of hospitals in the VA system, and the National Surgery Quality Improvement Program, which provides a valuable source of information for outcomes research studies. However, I have several questions regarding some of the conclusions.

Dr. Khuri states that surgical volume should not be used as surrogate measures of quality. I would like to know if you think it is possible from the analysis performed to refute the association of risk-adjusted mortality with provider volume that other studies have shown.

First, as you have noted, mortality is not the most appropriate outcome to measure for a number of procedures studied—for example, for cholecystectomies. As was noted, the procedures studied were not complex high-risk procedures and the relative homogeneity of the VA population I think limits the generalizability of the findings. Thus, I would question the comparability of these results to published works on complex high-risk surgery.

Also, a number of the reference studies did adjust for comorbidities and severity of illness, in contrast to your interpretation of these studies in the paper I had the opportunity to look at. Most significantly, I question whether it is possible to draw any conclusions regarding high-volume providers *versus* low-volume providers in your database, as you did not separately identify and analyze patients from any high-volume providers as a group. Were there enough high-volume providers out of the group of 125 hospitals to even consider this? I would like to have you describe the distribution of hospital volume by procedure to see if in fact a comparison between high- and low-volume providers would be possible.

Also, you looked at volume quartiles, but given the ranges in the quartiles, they really spend well in intermediate-volume ranges. So again this would not yield an appropriate comparison with high-volume providers to intermediate and low-volume providers.

You noted that the VA hospitals could be affiliated with academic medical centers and may have a crossover of high-volume surgeons performing procedures in low-volume VA hospitals adjacent to the academic medical centers. So I would like to know how you thought about adjusting for physician volume as well.

I am also interested to know how completely in-hospital and 30-day mortality were reported, and again note that the published studies that you have compared your results to looked at in-hospital mortality *versus* yours which is looking at 30-day mortality. So how completely was that reported? Was there missing data? If so, how was that handled?

Given some of the volume adjustment techniques that you used, were there any temporal trends in the data? Because you did some extrapolations for missing data across the study period.

Also, I think another area that I would like to hear your thoughts on is whether there are any regional effects across the VA system, given that there are well-documented regional variations in health care in the private sector that have been published and studied by Landberg and others.

Also, I would like to have you comment on the structural capabilities of the VA relative to the private sector and reconsider drawing the conclusion that there is not a difference between the VA system and the private sector with respect to the structure and process differences.

Last, you looked at risk factors at the procedure level. And it looks like some of the risk factors of interest were not considered in some of your regression models, such as age or history of TIA for risk of mortality for carotid endarterectomy. So I would like to know how you considered these clinical indicators with respect to the face validity of your regression models.

Also, you had some comparisons with the registry data. And I would like to know if you feel that you adequately risk-adjusted for those comparisons with the published studies that looked at the positive association with high-volume providers.

In conclusion, I think the main finding from your paper should be that for VA hospitals which are not high-volume hospitals, outcomes cannot be differentiated across low- and intermediate-volume facilities based on volume. Great caution should be exercised in generalizing these data to the private sector which includes high-volume providers. Hence, there remains great opportunity in the VA system to discern clinical indicators of quality and I think you have a tremendous database to further explore this.

DR. KHURI: We did not at all imply that these data are applicable to the private sector. They are specific to the VA, and, as such, may not be generalizable. The findings of this study, which was limited to prevalent operations of medium complexity, also may not be applicable to highly complex operations. The high-volume VA hospitals are not comparable to the high-volume hospitals in the private sector. The volume of surgery in the VA hospitals in our study is comparable to the low- and intermediate-volume hospitals in similar studies from the private sector. However, there was enough variation in volume between the hospitals in our study to allow for a meaningful comparison between quartiles of volume.

I agree with you that the NSQIP database cannot provide complete information on individual surgeons because most surgeons in the VA operate inside and outside the VA. The NSQIP database only accrues data on their VA patients. For this reason, we have not performed provider-specific analyses.

Thirty-day follow-up is conducted by the NSQIP nurse at each institution. Mortality is also verified by the benefits database in the VA which is highly reliable. Hence there were no missing mortality data in this study. We agree with you that risk-adjusted mortality alone might not be a good indicator of quality of care. However, we have just completed a study similar to this one in which we examined the relationship of volume of surgery to postoperative morbidity. Preliminary results from this study also fail to demonstrate an inverse relationship between volume of surgery and postoperative morbidity.

We did not observe temporal trends in the volume data and we did not observe regional variations between VA hospitals, although we did not specifically look for such variations. We did not conclude that "there is not a difference between the VA system and the private sector." On the contrary, we underscored these differences in the manuscript and indicated that our findings were specific to the VA and not generalizable to the private sector.

In terms of our predictive risk models, we constructed a separate model for each operation which took into account 67 preoperative risk factors, including age and neurologic history. All potentially important risk factors were analyzed. The c-indices of these models indicated a high level of predictability and all models had clinical face validity. We strongly believe that risk adjustment based on prospective collection of specific clinical data by dedicated nurses is more superior than risk adjustment based on codes contained in administrative databases.

DR. ERIC MUNOZ (Newark, New Jersey): I want to compliment the VA group for the excellent work they have been doing in health outcomes, and think that the discussion in the last 45 minutes has really gone to show the conflict in information that both government and payers have on this data.

First of all, I think it is common sense that for many procedures, the more you do the better you get. That has been shown over the years. The problem becomes when you look at the fact that most surgeons in most hospitals are low-volume and the fact that we

have some 3,000 hospitals and a couple hundred thousand physicians doing procedures, it gets very complicated to try to ask and answer the question: Which operations or which procedures should in fact be centralized?

I agree absolutely with Dr. Brennan that surgeons must take leadership in this because of the fact there are other groups that are very interested. I think papers such as yours and the work you have been doing is really key in this. And I would compliment the Association for their continued support of this. I think it is very, very important.

DR. JOHN L. CAMERON (Baltimore, Maryland): In most of our cost and outcome studies carried out at The Johns Hopkins Hospital we have looked at high-, mid-, and low-volume for surgeons or hospitals, and have compared the high volume to the low volume as well as the high to the mid, and the low to the mid. With your O/E ratios, it seems to me that you are always comparing to the mean, which should mute your results. If we only compared high volume to mid volume, and low volume to mid volume, which I think your O/E ratios do, then I think we would miss the difference that we see when we compare high to low. So unless I have read your manuscript incorrectly, I think you are comparing to the risk-adjusted mean, which would tend to mute your results.

I would also like to emphasize that a low-volume VA hospital could be right next door to a high-volume tertiary university hospital and the attendings in the low-volume VA could be very large-volume surgeons from the high-volume university hospital. And therefore, I question the VA System as a valid model to look at volume outcome studies.

DR. KHURI: These are very good thoughts, Dr. Cameron. The O/E ratio does not affect the comparison between the volumes of the hospitals *per se*. It is a risk-adjustment tool which only affects the outcome, *i.e.* the mortality rate. Since the O/E ratio is based on an expected outcome derived from a model based on all the hospitals performing a specific operation, it is not exactly a mean, and one need not necessarily have the same number of outliers on both ends of the range of the O/E ratio.

You may be absolutely right in your assertion that the VA model might not be the ideal one to investigate volume/outcome relationships. But one thing we have learned from this and other NSQIP studies which I think is fairly applicable to the private sector, is that quality of care is probably more determined by the processes and structures of a specific surgical service than by the volume of surgery performed by that service. This is also reflected in Dr. Harmon's presentation which preceded this one. Dr. Harmon showed us very clearly that low-volume surgeons working in high-volume hospitals had comparable outcomes to the high-volume surgeons working in these hospitals, the implication being that these high-volume hospitals had good processes and structures and, as such, were good hospitals. It is certainly true in the VA that it is the quality of the compendium of processes and structures in an institution which determines the outcome, not the volume of surgery performed.

DR. JEROME J. DECOSSE (New York, New York): Thank you Dr. Khuri, for a very interesting address. A comment and a question.

The comment parallels Dr. Cameron's remarks that with a very complex study such as you have with numerous categories, it is virtually impossible to avoid some error in modeling. And with

accumulation, you induce a regression to the mean and tend to favor the null hypothesis.

My question pertains to the mortality rate for colon surgery, which was 6.9%. It seemed rather high. You pointed out that 18% of these were emergency cases. Does that mortality reside within the 18%? And is there any difference in the distribution by hospitals?

DR. KHURI: First of all, in terms of your comment regarding the regression of the mean, I do agree with it, except that this is the reason why we used three other statistical analyses which did not depend on regression to the mean. These analyses also confirmed the results obtained with the regression analysis.

In terms of the mortality for colectomy, the 6.9% figure is fairly comparable to the 6.3% and the 6.0% figures recently reported in state registries and other large databases. All these are crude mortality rates and are not adjusted for the preoperative severity of illness. It is our impression, which we hope to validate in the future, that the VA patient population is generally sicker than the patient population at large. If so, this and the fact that 18% of the colectomy operations in our database were emergent, should explain the slightly higher mortality rate which we reported compared to those reported in published registries. We did observe a variation in the unadjusted and the risk-adjusted mortality rates for colectomy among the various hospitals, but we did not look into other details of this operation.